General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.
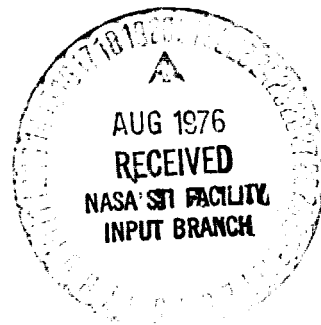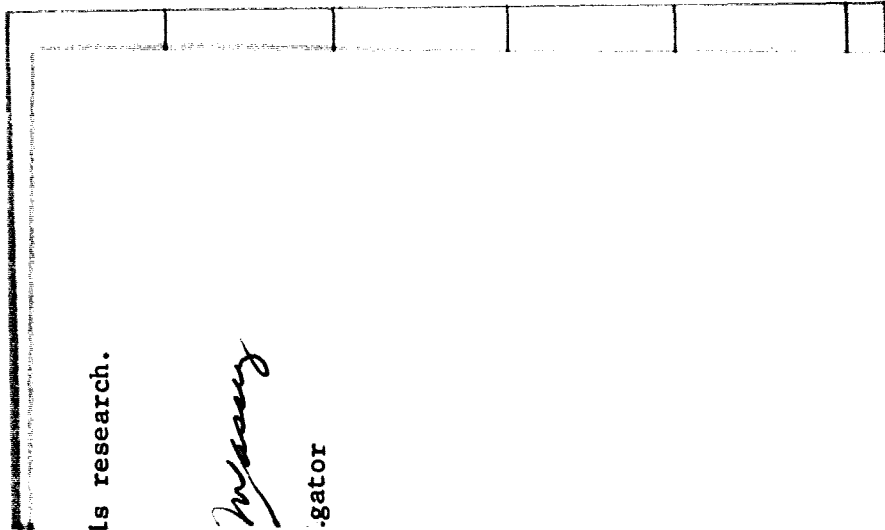
ls research.

gator

AUG 1976
RECEIVED
NASA STI FACILITY
INPUT BRANCH

Department of

ELECTRICAL ENGINEERING

UNIVERSITY OF NOTRE DAME, NOTRE DAME, INDIANA

August 17, 1976

SEMI-ANNUAL STATUS REPORT

COORDINATED DESIGN OF CODING
AND MODULATION SYSTEMS

(NASA Grant NSG 5025)

(Suppl. No. 2)

Submitted to: NASA Goddard Space Flight Center
Greenbelt, Md. 20771
  ATT: Dr. Thomas J. Lynch (Code 930)
         Mr. William B. Poland, Jr. (Code 730.4)
         Mr. John Rende (Code 814.1)

and

NASA Scientific and Technical Information Facility
P.O. Box 8757
Baltimore/Washington International Airport
Baltimore, Md. 21240

Principal Investigator: Dr. James L. Massey
Freimann Professor of Electrical Engineering
University of Notre Dame
Notre Dame, Indiana 46556

Research Assistants: T. Ancheta
D. Bordelon
G. Lauer
S. Narasimhan
J. Tomcik

RESEARCH PERIOD REPORTED: December 1, 1975 to May 31, 1976.

I. INTRODUCTION

In the following sections, we report briefly on the progress made in each of the indicated areas during the period December 1, 1975 to May 31, 1976, as well as on the plans, if any, to pursue further research in that area.

II. CONVOLUTIONAL CODES

A. Partial-Unit-Memory Codes

In previous research under this grant, L. Lee developed the class of unit-memory (UM) convolutional codes [1]. The UM codes are byte-oriented, and hence are natural choices for use as inner codes in concatenated coding systems having a Reed-Solomon outer code [2].

In a UM code, the encoded branch of n_o bits at time t may be written

$$\underline{x}_t = \underline{u}_t G_0 + \underline{u}_{t-1} G_1 \tag{1}$$

where $\underline{u}_t$ is the k_o bit information byte at time t, and where G_0 and G_1 are $k_o \times n_o$ binary matrices. Implementation of (1) requires an encoder with at most k_o binary delay cells, namely those to store the components of $\underline{u}_{t-1}$. [For a general convolutional code with memory M, where on the right in (1) there would be the further terms $\underline{u}_{t-2} G_2 + \ldots + \underline{u}_{t-M} G_M$, at most Mk_o binary delay cells would be required.] This led Lee to define the state-complexity [which hereafter we shall refer to as the virtual state-complexity] as k_o for UM codes [or, for general codes, as MK_o].

Lee showed that, for a given virtual state-complexity, the maximum free distance (d_∞) over the class of all convolutional codes is achieved within the class of unit-memory codes, and he found several cases of practical interest where the UM codes were strictly better than conventional

convolutional codes [i.e., codes where $\gcd(k_o, n_o) = 1$] with the same rate $R = k_o/n_o$ and the same virtual state-complexity [1]. This is important because virtual state-complexity, μ_V, implies the upper bound 2^{μ_V} on the number of states in the corresponding Viterbi decoder, a bound met with equality by all good conventional convolutional codes that have been proposed for use with Viterbi decoding.

It was realized by the principal investigator that the <u>true</u> state-complexity, μ, of a unit-memory code is

$$\mu = \text{rank } (G_1) \tag{2}$$

as can be seen from the following argument. Because μ is the rank of G_1, there exists a <u>non-singular</u> $k_o \times k_o$ binary matrix T such that the matrix $G_1' = TG_1$ has only zeroes in its first $k_o - \mu$ rows and has μ linearly independent vectors in its last μ rows. Thus, the new unit-memory code, defined by $G_0' = TG$ and G_1', is completely equivalent to the former code (same d_{free}, same catastrophic properties, etc.), but its encoding rule

$$\underline{x}_t = \underline{u}_t \, G_0' + \underline{u}_{t-1} \, G_1' \tag{3}$$

can be implemented using only μ delay cells since there is no need to store the first $k_o - \mu$ components of $\underline{u}_{t-1}$. For obvious reasons then, we call a UM code with $\mu < k_o$ a <u>partial-unit-memory</u> (PUM) code. Without loss of essential generality, we shall always take G_1 for a PUM code to be a matrix whose first $k_o - \mu$ rows are all zeroes.

The principal investigator's interest in PUM codes began with his noting that Lee's ($n_o = 8$, $k_o = 4$) UM code with $\mu_V = Mk_o = 4$ [which has $d_\infty = 8$] had

$$G_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

which has rank only $\mu = 3$ [as can be noted most easily from observing that

its rows sum to the all-zero vector.] The transformation

$$T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

then gives the $\mu = 3$ PUM code with

$$G_1' = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

This code also has $d_\infty = 8$ but requires only $2^\mu = 8$ states in its Viterbi decoder; this should be compared to $2^{\mu_V} = 16$ states for the parent UM code and $2^{Mk_o} = 32$ states in the Viterbi decoder for the best rate $R = k_o/n_o = 1/2$ conventional code. This four-fold reduction in the number of states in the Viterbi decoder, between the PUM code and the best conventional code with the same $d_\infty = 8$, is graphic illustration of the practical potential of PUM codes as inner codes in concatenated coding systems.

A detailed study of PUM codes has been undertaken by G. Lauer as his master's dissertation research under the direction of the principal investigator. This research is fully described in:

G. S. Lauer, "Some Optimum Partial-Unit-Memory Convolutional Codes," Tech. Rpt. No. EE-767, Dept. of Elec. Engr., Univ. of Notre Dame, Notre Dame, Ind., July 28, 1976.

In this report, Lauer gives PUM codes with $\mu = k_o - 1$ having the same d_∞ as the best UM codes (but requiring only half as many decoder states) for the following cases: $R = 1/4$ and $\mu = 1,2,3$ and 4; $R = 1/3$ and $\mu = 1$ and 4; $R = 1/2$ and $\mu = 1$ and 3; and $R = 2/3$ and $\mu = 1$ and 2.

As Lauer's new PUM codes appear to cover all cases of practical interest, no further searching for PUM codes is planned.

B. Interleaving Requirements for Unit-Memory Codes

For a concatenated coding system using an RS outer code together with an inner convolutional code which is decoded by a Viterbi decoder, it is necessary for good system performance that the inner code be interleaved sufficiently so that, in each interleaved stream, the decoding errors of the Viterbi decoder appear to be independent [2]. Normally, this requires at least $\Delta+1$ interleaved streams where Δ is the decoding delay [2].

Some preliminary experimentation with simple unit-memory (UM) codes has indicated that a decoding delay of only $\Delta = 1$ may suffice to give virtually the same decoding error probability as the theoretically optimum $\Delta = \infty$. This would be of considerable practical importance in minimizing the interleaving requirements for the inner code of a concatenated coding system. This possibility is presently under intensive investigation by J. Tomcik, and it is expected that his research will be completed and reported in the next half-year.

III. PHASE-TRACKING ERROR EFFECTS ON CODING PERFORMANCE

One of the primary research goals for this grant has been the determination of the effect of phase-lock loop (PLL) tracking error on the coding system performance. The keystone of our research into this topic has been usage of the channel <u>cut-off rate</u>, R_o, as the measure of quality of a modulation system to be used with coding [3].

An intensive investigation of the effect on R_o due to phase-tracking-error in the PLL is presently being conducted by S. Narasimhan. His preliminary work indicates that, in the E/N_o range of practical interest, this PLL tracking error places a penalty of about 2 db on R_o, but that nearly all of this loss can be recovered by utilizing the receiver's estimate of the

tracking error. It is expected that this work will have progressed sufficiently so that a substantial initial report can be issued in the next half-year.

IV. OPTIMUM MODULATION SIGNAL SETS

Again with R_o being used as the criterion of goodness for the modulation system, an extensive investigation has been made to find optimum modulation signal sets for a <u>non-white</u> Gaussian noise channel. The non-white model was chosen to admit applicability of the research to real channels with memory effects such as intersymbol interference.

This research was undertaken by D. Bordelon as his master's dissertation research under the direction of the principal investigator. His results are fully reported in:

D. L. Bordelon, "On Signal Design by the R_o Criterion for Non-White Gaussian Noise Channels," Tech. Rpt. No. EE-768, Dept. of Elec. Engr., Univ. of Notre Dame, Notre Dame, Ind., July 28, 1976.

Perhaps the most interesting product of Bordelon's research was an heuristic selection rule for signal sets based on a "water-filling" argument. With infinitely many signals in the signal set and a Gaussian distribution on their usage, Bordelon found that the optimum (in the sense of maximizing R_o) distribution of signal energy along each coordinate axis had a "water-filling interpretation," similar to, but not coinciding with, the water-filling interpretation for the energy distribution when the signal set achieves channel capacity [4]. Then, for a given finite number M of signals, Bordelon found that, when he selected the signals so as to yield the same energy along each coordinate axis as determined by the water-filling rule, the M-ary signal set was virtually optimum. Thus, Bordelon's heuristic signal selection

rule provides a simple method to select a "good" signal set for channels with non-white Gaussian noise.

No further research in this area is planned for the next half-year.

V. DATA COMPRESSION

Part of the research under this grant has been concerned with the use of error-correcting codes to perform "data compression" by the technique of syndrome source coding. This work is being conducted by T. Ancheta who, already under this grant, has developed an important practical extension [5] of syndrome source coding which he termed "noiseless universal syndrome source coding" (NUSSC).

During this research period, Ancheta has developed strong bounds on the ultimate performance attainable with syndrome source coding. Letting $R(D)$ be the rate-distortion function for the (asymmetric) binary memoryless source (BMS), Ancheta has shown that <u>linear source encoding</u> (of which syndrome source coding can be considered a special case) has an ultimate performance $R_L(D)$ such that $R_L(D) \geq R(D)$ with equality when and only when $D = 0$ or $D = p$ (where p is the probability that the BMS emits a 1.) Moreover, for any syndrome source coding scheme based on a code with partially-symmetric information sets (PSIS), Ancheta has proved that an even stronger lower bound applies. [A code is said to have the PSIS property if there is a subset of its information sets (an information set is any set of k positions in which all 2^k codewords are distinct) which contains each position the same number of times. This is a weak constraint satisfied by most codes, for instance, by all cyclic codes.] These results imply a dim future for syndrome source coding in situations where non-zero distortion in the reconstructed source output is acceptable. Fortunately for the future of syndrome source coding, the practical case of greatest interest is zero distoriton (i.e., "noiseless

coding") for which syndrome source coding gives the same ultimate performance as any form of source coding.

Ancheta has also, during the past half year, been testing the effectiveness of NUSSC on real telemetry data obtained from the Goddard Space Flight Center. Preliminary results have been highly encouraging.

It is expected that Ancheta will complete his investigation of data compression by syndrome source coding during the next half year and that a detailed report thereon will be issued.

The principal investigator has developed a very simple data compression scheme, closely related to NUSSC and to Lynch-Davisson-Schalkwijk-Cover (LDSC) noiseless universal coding (cf. [5]), which may have practical applicability. We shall call this scheme the weight-and-error-locations (WEL) method of data compression.

The WEL scheme operates as follows. Let m and r be any two given positive integers with $r \leq m$. Let $\underline{e} = [e_1, e_2, \ldots e_n]$ be $n = 2^m$ output bits from the source which, as the notation suggests, we are treating as an "error-pattern." If there are t non-zero bits in $\underline{e}$ (i.e., t errors) where $t < 2^r - 1$, then $\underline{e}$ is encoded as $\underline{w} * \underline{p}_1 * \ldots * \underline{p}_t$ (* denotes concatenation) where $\underline{w}$ is the m-bit radix-two form of t, and where (if the errors are in positions $i_1 < i_2 < \ldots < i_t$) $\underline{p}_1$ is the m-bit radix-two form of $i_1 - 1$ and $\underline{p}_j$, for $2 \leq j \leq t$, is the m-bit radix-two form of $i_j - i_{j-1} - 1$. When $t \geq 2^r - 1$, then $\underline{e}$ is encoded simply as $\underline{1} * \underline{e}$ where $\underline{1}$ denotes the all-one binary m-tuple. Note that both the source encoding and decoding could be accomplished in an obvious way using only two simple counters.

<u>Example</u>:   With m = 4 (so that n = 16) and r = 2, the WEL scheme would

operate as follows:

Weight of $\underline{e}$	Form of encoded word	Length of encoded word	Compression factor
0	[0 0]	2	8
1	[0 1] * $\underline{p}_1$	6	2.67
2	[1 0] * $\underline{p}_1$ * $\underline{p}_2$	10	1.60
$\geq$ 3	[1 1] * $\underline{e}$	18	0.89

When this scheme is used to encode a BMS with p = .0312, the attained

compression ratio is

$$\gamma = \frac{16}{\text{Average Length of Encoded Word}}$$

$$= \frac{16}{2(1-p)^{16}+6\binom{16}{1}p(1-p)^{15}+10\binom{16}{2}p^2(1-p)^{14}+18[1-(1-p)^{16}-\binom{16}{1}p(1-p)^{15}-\binom{16}{2}p^2(1-p)^{1}}$$

$$= \frac{16}{4.04} = 3.96$$

which should be compared to the theoretical limit $\gamma = 1/H(p) = 5$ for any

noiseless source coding scheme for this particular BMS.

The WEL data compression scheme with r = m is very closely related to

LDSC coding. The only difference is that, in the LDSC scheme, after sending

the same heading $\underline{w}$ as in the WEL scheme, one sends the "index" of the error

pattern in a list of all $\binom{n}{w}$ n-tuples of weight w rather than simply the w

locations of the one's. The LDSC scheme thus minimizes the number of bits

used to identify the particular n-tuple $\underline{e}$ of weight w at the cost, however,

of a considerably more complicated coding rule.

The WEL scheme is also closely related to NUSSC in which $\underline{e}$ is coded as

$\underline{w}$ * $\underline{s}$ where $\underline{s}$ is the syndrome for $\underline{e}$ in a particular t-error-correcting code

[5]. Thus, WEL differs from NUSSC in that the former specifies $\underline{e}$ by the

location of its t one's whereas the latter specifies $\underline{e}$ by its syndrome with

respect to some t-error-correcting code. It is interesting to note that the mt digits used to specify the t error locations in WEL are precisely the number of syndrome digits specified by the usual bound [6] for a t-error-correcting BCH code of length $2^m - 1 = n-1$. Thus, the efficiencies of WEL coding and NUSSC will generally be very nearly the same for the cases of practical interest where this bound mt for the number of syndrome bits in the BCH code is either exact or very tight. Moreover, the WEL scheme appears even simpler to implement than NUSSC.

REFERENCES

[1] L.-N. Lee, "Short Unit-Memory Byte-Oriented Binary Convolutional Codes Having Maximal Free Distance," IEEE Trans. Info. Th., Vol. IT-22, pp. 349-352, May 1976.

[2] L.-N. Lee, "Concatenated Coding Systems Employing Unit-Memory Convolutional Codes and Symbol-Oriented Optimal Decoding Algorithms," Appendix in Final Report, First Year, "Coordinated Design of Coding and Modulation Systems," NASA Grant NSG 5025, Dept. of Elec. Engr., Univ. of Notre Dame, Notre Dame, Ind., June 14, 1976.

[3] J. L. Massey, "Coding and Modulation in Digital Communications," in Proc. Int. Zurich Seminar on Digital Comm., pp. E2(1)-E2(4), March 12-15, 1976.

[4] R. G. Gallager, Information Theory and Reliable Communication. New York: Wiley, 1968, pp. 344-345.

[5] T. C. Ancheta, Jr., "Syndrome-Source-Coding and Its Universal Generalization," IEEE Trans. Info. Th., Vol. IT-22, pp. 432-436, July 1976.

[6] W. W. Peterson, Error-Correcting Codes. Cambridge, Mass.: M.I.T. Press, 1961, p. 165.